

# Forecasting Generalized Quantiles of Electricity Demand: A Functional Data Approach

Brenda López Cabrera\*  
Franziska Schulz\*

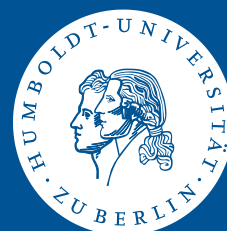


\* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



# Forecasting Generalized Quantiles of Electricity Demand: A Functional Data Approach. <sup>\*</sup>

Brenda López Cabrera, Franziska Schulz<sup>†</sup>

May 30, 2014

## Abstract

Electricity load forecasts are an integral part of many decision-making processes in the electricity market. However, most literature on electricity load forecasting concentrates on deterministic forecasts, neglecting possibly important information about uncertainty. A more complete picture of future demand can be obtained by using distributional forecasts, allowing for a more efficient decision-making. A predictive density can be fully characterized by tail measures such as quantiles and expectiles. Furthermore, interest often lies in the accurate estimation of tail events rather than in the mean or median. We propose a new methodology to obtain probabilistic forecasts of electricity load, that is based on functional data analysis of generalized quantile curves. The core of the methodology is dimension reduction based on functional principal components of tail curves with dependence structure. The approach has several advantages, such as flexible inclusion of explanatory variables including meteorological forecasts and no distributional assumptions. The methodology is applied to load data from a transmission system operator (TSO) and a balancing unit in Germany. Our forecast method is evaluated against other models including the TSO forecast model. It outperforms them in terms of mean absolute percentage error (MAPE) and achieves a MAPE of 2.7% for the TSO.

Keywords: Electricity, Load forecasting, FPCA

JEL classification: G19, G29, G22, Q14, Q49, Q59

## 1. INTRODUCTION

With the liberalization of energy markets, the amount of risk borne by operators and market participants has increased substantially. Statistical tools can be beneficial to

---

<sup>\*</sup>The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 “Ökonomisches Risiko”, Humboldt-Universität zu Berlin is gratefully acknowledged.

<sup>†</sup>Ladislaus von Bortkiewicz Chair of Statistics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email:lopezcab@wiwi.hu-berlin.de,sulzfran@hu-berlin.de

assess and manage their risk. Amongst energy, electricity stands out due to its limited storability. Supply and demand have to be balanced out at every point in time. Since electricity is mainly traded in a day-ahead market, short term adjustments in supply due to forecasting errors can lead to enormous financial losses. Therefore, load forecasting is extremely important for energy suppliers, transmission system operators, financial institutions and other participants in electricity market and a crucial process in the planning and operation of electric utilities.

There is a vast literature on how to forecast electricity load, most of them concentrating on deterministic forecasts. For an overview on common methods see Weron (2007) or Taylor and McSharry (2007). However, for a sustainable risk management of utility operators not only a forecast of expected demand, but also knowledge about the uncertainty and dispersion of future load plays an important role. This points towards the use of probabilistic forecasts. While in different areas of forecasting such as macro-economics and finance (Tay and Wallis 2000), meteorology (Leutbecher and Palmer 2008) or renewable energy production (Bremnes 2004; Pinson et al. 2007) probabilistic forecasts are already well established, there is a lack in literature on probabilistic forecasts for electricity demand. A notable exception is the work by Hyndman and Fan (2010) which uses a mixture of temperature simulation, economic scenarios, and residual bootstrapping to obtain long-term density forecasts of electricity demand. Others like Cottet and Smith (2003) use Bayesian modeling in a multi-equation regression model to forecast intraday electricity load and briefly discuss model averaging for probabilistic forecasts.

Short-term probabilistic forecasts yield important information for utility operators for decisions e.g. on purchasing and generating electricity and load scheduling. They are crucial for risk management and can be used to derive risk measures such as probability of exceedance levels (Taylor 2008; Bellini et al. 2014). In this article we propose a methodology to obtain probabilistic forecasts by employing functional data analysis of generalized quantile curves.

With generalized quantiles we refer to quantiles (Koenker and Bassett Jr 1978) and expectiles (Newey and Powell 1987). Both are tail measures and uniquely characterize the conditional distribution of a random variable. Furthermore, for a large class of decision-making problems, optimal solutions correspond to quantiles of a conditional predictive distribution (Gneiting 2011). In fact, in a wide range of fields including weather events, extreme natural hazards, genomics, risk management, energy demand and portfolio allocation among others, tail indices provide useful information that goes beyond the mean and median. These tail indices constitute curves that can be treated in a functional principal component analysis (FPCA) context.

The idea of FPCA is to identify the main risk drivers by a small number of factors combined with random noise. The resulting factors are then to be (cor)related with exogenous variables, which will allow us to study phenomena contingent to extreme risks. Functional data analysis (see Ramsay and Silverman (2005)) is an extension of multivariate data analysis to functional data. There are few studies that apply methods from functional data analysis to load forecasting, amongst which are Shang (2013), Goia et al. (2010) and Antoch et al. (2008). Others like Cho et al. (2013) reduce dimension using a hybrid approach that combines a generalized additive

model for the weekly averages of the load and curve linear regression models for the dependence structure across consecutive daily loads.

Two recent studies on functional data analysis of tail events are Guo et al. (2013), who do the dimension reduction with weighted  $L_1$  and  $L_2$  norm, where the weights are sign sensitive, and Tran et al. (2014) who develop an analogue of PCA of tail curves in an asymmetric norm. However, both studies rely on independence between functional observations. In many fields however, such as energy demand modeling, the dependency between curves needs to be taken into account at the core of the modeling. In our study we allow for temporal dependence between functional observations and refer to results from Hörmann and Kokoszka (2010). The dependence between tail curves can be exploited for forecasting, which provides useful information to support modeling, pricing and trading. Our approach has several advantages: It allows for flexible inclusion of explanatory variables and does not require distributional assumptions for the tails curves. Furthermore, treating load curves as functional data has the advantage that one step ahead forecasts yield forecasts for the whole next day and that forecasts are continuous functions and thus available for every point in time. We expect that exogenous variables like meteorological factors do not only effect the amount of electricity consumed, but also the distribution over the day and thus the shape of the load curves (Engel et al. 1986; Harvey and Koopman 1993; Taylor and McSharry 2007).

We illustrate our approach with data on quarter-hourly electricity consumption of a transmission system operator (TSO) and a balancing unit (BU) in western Germany. Variations in the intradaily pattern are explained using weather variables and meteorological forecasts. The proposed model is shown to perform better than well-known methods, such as a deterministic similar-day approach, the Holt-Winter Exponential smoothing and the forecast provided by TSO Amprion. It achieves on average 2.7% mean absolute percentage error (MAPE) in the one-day forecasting period for the TSO.

Our article is structured as follows. Section 2 gives a brief introduction of the load data. In Section 3 we define FPCA of generalized quantiles that will be used to produce probabilistic load forecasts, together with its estimation algorithm. Section 4 discusses the modeling and estimation of the electricity demand data. Section 5 describes the forecast performance with respect to other methods. Section 6 concludes the paper. All computations in this paper were carried out in R. The electricity load data and forecast electricity load data was obtained from TSO Amprion and the balancing unit Stadtwerke Saarbrücken. The data source for the temperature data is Deutscher Wetterdienst (DWD); for the meteorological weather forecast data the data source is WeatherOnline. We thank Dr. Ulrich Römer and Herrad Werner for providing us the data. To simplify notation, in the following dates are denoted with `yyyymmdd` format. Supplementary materials for this article are available online.

## 2. ELECTRICITY DEMAND DATA

The German electricity market, which was liberalized in 1998, is Europe’s largest, with annual power consumption of around 500 TWh and a generation capacity of 125 GW (Eurostat 2014). The four German TSOs (Amprion, Tennet TSO, 50Hertz

Transmission and TransnetBW) are responsible for maintaining a stable and reliable system and to maintain balance between electricity generation and consumption. All market participants are organized in balancing units (BU). Each BU has a BU manager who is responsible for the balance within the unit. Electricity is traded mainly in the day-ahead market, which closes at 12pm. Before, each BU manager has to submit a load schedule to the corresponding TSO specifying the expected load for each quarter hour of the next day. Deviations from the specified load can still be adjusted in the intra-day market. The intra-day market is a continuous market where contracts can be traded until 45 minutes before delivery. However, the intra-day market is less liquid than the day-ahead market and therefore neglected in this study. The TSOs balance out differences between the forecasted load of the BU and actual consumption in order to ensure a stable system. Precise forecasts of the area's consumption are essential in order to have sufficient capacity available. For deviations between forecasted and actual load BUs have to pay a price which usually greatly exceeds the price at the spot market. Therefore, for BU managers improvements in their forecasting performance directly leads to enormous cost reductions.

For the empirical work of this article, we use electricity demand data of the TSO Amprion and the BU Stadtwerke Saarbrücken. Both datasets are freely accessible on their websites. The TSO Amprion operates in the west of Germany. The BU is located within the balancing area of Amprion. The analysis is based on quarter hourly data of electricity consumption from 20100101 to 20123112. Summary statistics are given in Table (1). The first two years of the data are used for in-sample fitting and the third year for an out-of-sample forecasting evaluation, given in Section 5. Figure (1) displays the two load data sets from 2010 to 2012. It is clearly visible that electricity consumption exhibits seasonal features over time and at different times of the day.

	Median	Mean	SD	Min	Max
TSO	22020	22050	4054.35	11850	34870
BU	402	401	95.71	152	630

Table 1: Summary Statistics of the load data (in MW)

It is a stylized fact that it contains yearly, weekly and intraday seasonal cycles and is sensitive to temperature changes (Engel et al. 1986; Taylor and Buizza 2002). During winter electricity consumption in Germany is higher than during the summer. Additionally, at weekends electricity consumption is usually lower than during the week. The typical intraday load profile shows a peak around noon, followed by valley in the afternoon and another peak in the evening at around 7pm. These seasonal patterns are quite predictable and therefore usually modeled deterministically. We express the observed load  $\tilde{Y}_s$  as

$$\tilde{Y}_s = \Lambda_s + Y_s, \quad s = 1, \dots, S, \quad (1)$$

where  $\Lambda_s$  is a deterministic seasonal component and  $Y_s$  is a stochastic component. We estimate the deterministic seasonal component separately for every quarter hour

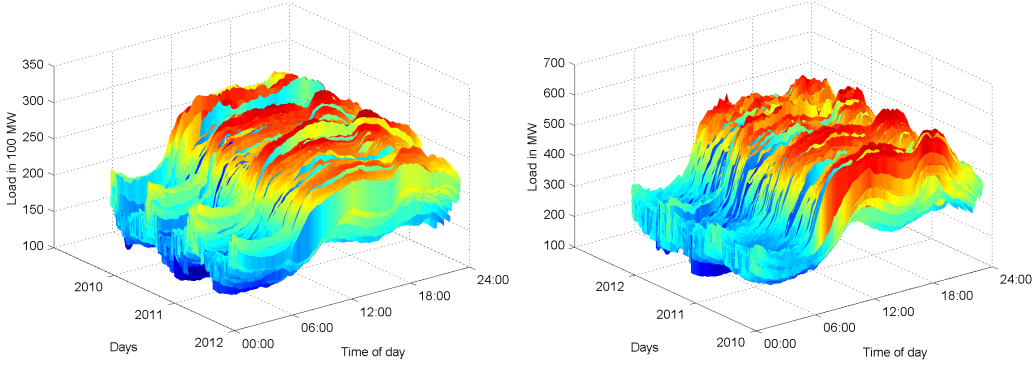


Figure 1: Electricity load curves of TSO - Amprion (left) and BU - Stadtwerke Saarbrücken (right) from 20100101 to 20123112.

of a day. It is specified as

$$\Lambda_{t,k} = a_t + b_t \cdot k + c_{1,t} \sin\left(\frac{2\pi k}{365}\right) + c_{2,t} \cos\left(\frac{2\pi k}{365}\right) + \sum_{i=1}^7 d_{i,t} \cdot D_{i,k}, \quad (2)$$

where  $t = 1, \dots, 96$  denotes the quarter hours of a day and  $k = 1, \dots, K$  the day, such that  $t \cdot k = s$ . The parameters  $a_t, b_t, c_{1,t}, c_{2,t}$  and  $d_{i,t}$  are estimated by ordinary least square regression.  $D_{i,k}$  is a set of dummy variables consisting of six dummies for the weekdays and one dummy for public holidays. They capture weekly seasonal behavior, while the sine and cosine functions capture yearly seasonalities. This approach is very close to the so called similar-day approach, which is a commonly used approach in industry to model and forecast electricity load.

As covariates for load modeling we include average daily temperature and hours of sunshine. The time series of both variables are displayed in Figure (??). For Saarbrücken temperature and hours of sunshine are measures from a weather station in Saarbrücken. For the TSO the average measures of three stations located in their area are taken. For forecasting we use meteorological day-ahead forecasts of the covariates, which are provided by WeatherOnline. For our analysis all covariates are deseasonalized.

### 3. METHODOLOGY

We propose a methodology that combines methods from generalized quantile regression and functional data analysis (FDA) in order to obtain probabilistic forecasts of electricity demand. FDA has gained importance with the advances in storing large sets of multivariate data. It has been applied in various fields of research ranging from bioscience to medicine and econometrics. For an overview on applications of FDA we refer to Ramsay and Silverman (2002) or Ferraty and Vieu (2006).

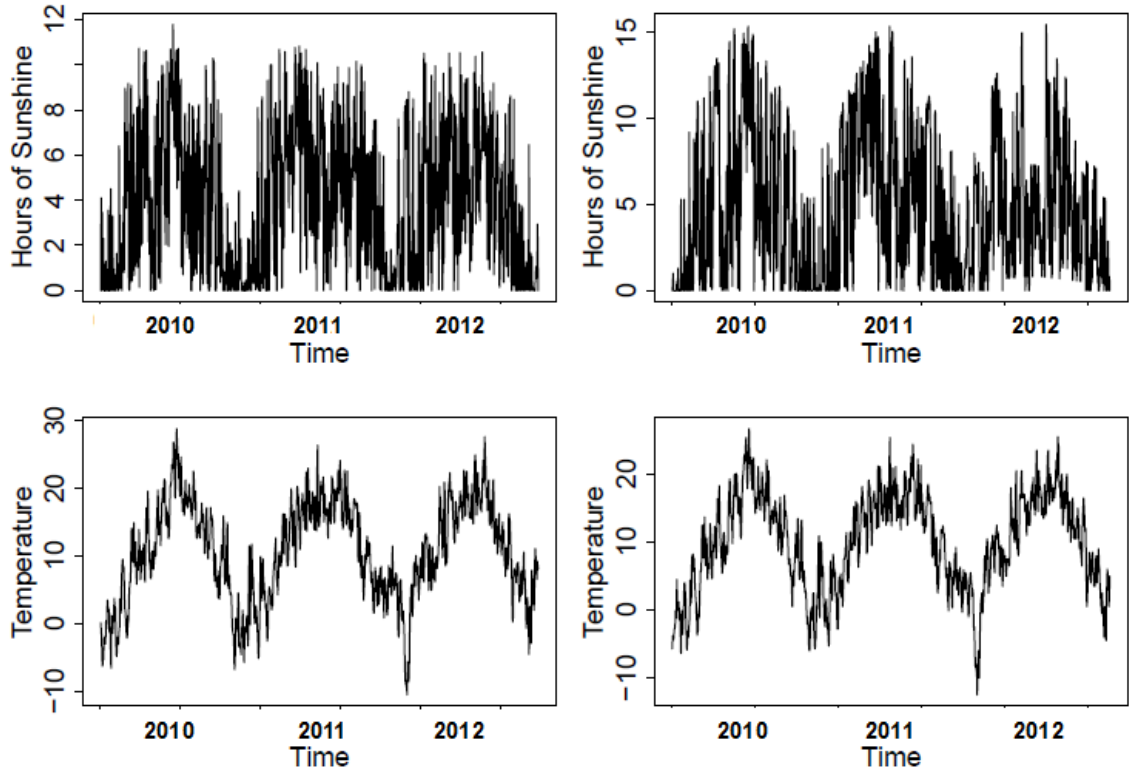


Figure 2: Average daily temperature and hours of sunshine for the area of the TSO Amprion (left) and Saarbrücken (right)

### 3.1 Generalized Quantiles

The distribution of a random variable  $Y$  can be characterized by its cdf  $F_Y(y)$ . The quantile functions of  $Y$  are defined as

$$Q_Y(\tau) = F_Y^{-1}(\tau) = \inf\{y : F(y) \leq \tau\}, \quad \tau \in (0, 1). \quad (3)$$

Like the cdf, the quantile function provides a full characterization of the random variable  $Y$ . For each  $\tau \in (0, 1)$  the quantile function can be formulated as the solution of a minimization problem:

$$Q_Y(\tau) = \arg \min_y E\{\rho_\tau(Y - y)\}, \quad (4)$$

where  $\rho_\tau(\cdot)$  is a loss function defined as

$$\rho_\tau(u) = u\{\tau - I(u < 0)\}, \quad (5)$$

which is in general asymmetric (Koenker 2005). A special case is the median, which corresponds to  $\tau = 0.5$ . The quantile function conditional on a (one-dimensional) covariate  $X$  is given by

$$Q_{Y|X}(\tau) = \arg \min_{f \in \mathcal{F}} E\{\rho_\tau(Y - f(X))\}, \quad (6)$$

where  $f(\cdot)$  is a nonparametric function of the covariate  $X$  from a set of functions  $\mathcal{F}$ , such that the expectation is well defined. Closely related to quantiles are expectiles introduced by Newey and Powell (1987), which can be obtained by a generalization of the loss function  $\rho_\tau(\cdot)$ :

$$l_\tau(X) = \arg \min_{f \in \mathcal{F}} E\{\rho_\tau^\alpha(Y - f(X))\}, \quad \alpha \in \{1, 2\} \quad (7)$$

$$\rho_\tau^\alpha(u) = |u|^\alpha |\tau - I(u < 0)|. \quad (8)$$

We call the solution of (7) a generalized quantile function. For  $\alpha = 1$  we obtain the loss function in (5) and the solution to (7) is a conditional quantile function. For  $\alpha = 2$  the solution to (7) is a conditional expectile function. Note that the conditional expectile corresponding to  $\tau = 0.5$  is the expected value  $E(Y|X)$ . Like quantiles, expectiles characterize the distribution of a random variable. While they are less intuitive to interpret than quantiles, they have the advantage of a better computational efficiency (Newey and Powell 1987) and are a coherent risk measure (Bellini et al. 2014).

If one is not only interested in estimating a single generalized quantile, but in estimating a collection of generalized quantiles, one faces the problem of crossing quantiles/expectiles. This is unfavorable, since it is theoretically impossible. We use an algorithm proposed by Schnabel (2011), that estimates the generalized quantile functions  $l_\tau(X)$  jointly as a surface on the domain of the independent variable  $X$  and the asymmetry parameter  $\tau$ , called expectile sheet. The algorithm is based on a least asymmetrically weighted squares (LAWS) criterion combined with P-splines. The expectile sheet can be expressed as

$$\mu(X, \tau) = \sum_{k=1}^K \sum_{l=1}^L a_{kl} B_k(X) \tilde{B}_l(\tau), \quad (9)$$

where  $B$  is a B-spline basis for the covariate  $X$ ,  $\tilde{B}$  is a B-spline basis for the asymmetry parameter  $\tau$  and  $A = [a_{kl}]$  is a matrix of coefficients. It is estimated by minimizing

$$\sum_{i=1}^n \sum_{j=1}^J w_i(\tau_j) \{Y_i - \mu(X_i, \tau_j)\}^2 + P, \quad (10)$$

where  $w_i(\tau)$  is a weight function and  $\mu(X, \tau)$  is the expectile sheet.  $P$  is a penalty term defined as  $P = \lambda_x \|DA\|_F + \lambda_\tau \|A\tilde{D}\|_F$ , where  $\|\cdot\|_F$  is the Frobenius norm and  $D$  and  $\tilde{D}$  are second order difference matrices. The penalty term controls the smoothness of the estimates and  $\lambda_x$  and  $\lambda_\tau$  are smoothing parameters. The algorithm was originally developed for expectiles and the weight function  $w_i(\tau)$  is given by

$$w_i(\tau_j) = \begin{cases} \tau_j & \text{if } Y_i > \mu(X_i, \tau_j) \\ 1 - \tau_j & \text{if } Y_i \leq \mu(X_i, \tau_j). \end{cases} \quad (11)$$

With a slight modification the algorithm can be adapted for quantiles as well. Following Schnabel and Eilers (2013) the modified weight function is given by

$$w_i(\tau_j) = \begin{cases} \frac{\tau_j}{\sqrt{\{Y_i - \mu(X_i, \tau_j)\}^2 + \delta^2}} & \text{if } Y_i > \mu(X_i, \tau_j) \\ \frac{1 - \tau_j}{\sqrt{\{Y_i - \mu(X_i, \tau_j)\}^2 + \delta^2}} & \text{if } Y_i \leq \mu(X_i, \tau_j), \end{cases} \quad (12)$$



where  $\delta$  is a small constant which is used to avoid numerical problems. In the following the covariate  $X$  used is time of the day and will be denoted by  $t$ .

### 3.2 Time series of functional data

Electricity demand is recorded in sequential form and shows a similar pattern each day. Naturally, metered demand can be divided into time intervals of one day. We are interested in generalized quantiles of metered intra-daily electricity demand. The generalized quantile functions defined in the previous section can be treated as realizations  $l_{\tau,k}(t)$  of a functional time series  $(l_{\tau,k}, k \in \mathbb{Z})$  defined on a compact set  $\mathcal{T}$  and for a fixed  $\tau \in (0, 1)$ . For notational convenience we suppress  $\tau$  in the following, i.e.  $l_{\tau,k}(t) := l_k(t)$ . Under stationarity  $l_k(t)$  have a common mean function  $E\{l(t)\} = \mu(t)$  and a common covariance function  $C(s, t) = \text{Cov}\{l(s), l(t)\}$  with  $s, t \in \mathcal{T}$ . Note that functional observations are intrinsically infinite dimensional. A common tool to reduce dimensionality is functional principal component analysis (FPCA). For a survey on FPCA we refer the reader to Shang (2014).

FPCA yields the directions of largest variability in the data and expresses the data as a weighted sum of the orthogonal principal component functions. If  $\int_{\mathcal{T}} C(t, t) dt < \infty$ , the covariance function induces the Kernel operator  $K : \phi \mapsto K\phi$  defined by  $(K\phi)(s) = \int_{\mathcal{T}} C(s, t)\phi(t)dt$ . Then,  $C$  has the representation

$$C(s, t) = \sum_{i=0}^{\infty} \lambda_i \phi_i(s) \phi_i(t), \quad (13)$$

where  $\phi_i$  for  $i = 1, 2, \dots$  are the orthogonal eigenfunctions and  $\lambda_i$  the corresponding non-increasing and non-negative sequence of eigenvalues of the operator  $K$ . The eigenfunctions are also called principal component functions. The principal component scores  $\alpha_i, i = 1, 2, \dots$  are given by  $\langle l(t), \phi_i \rangle$ , where  $\langle \cdot, \cdot \rangle$  denotes the inner product. That is, the principal component scores are the projection of  $l(t)$  in the direction of the corresponding principal component, with  $E(\alpha_i) = 0$  and  $\text{Var}(\alpha_i) = \lambda_i$ . Using the Karhuhen-Loève expansion we can express each function  $l_k(t)$  as

$$l_k(t) = \mu(t) + \sum_{i=1}^{\infty} \alpha_{ik} \phi_i(t). \quad (14)$$

The functions  $l_k(t)$  can be approximated by a finite sum of the first  $m$  principal components, called truncated Karhuhen-Loève expansion:

$$l_k(t) \approx \mu(t) + \sum_{i=1}^m \alpha_{ik} \phi_i(t) \quad (15)$$

Note that only the principal component scores are varying over time, whereas the mean and principal component functions are time invariant. This observation is exploited for forecasting.

### 3.3 Estimating FPCA for generalized quantiles

In practice the mean, eigenvalues and eigenfunctions are unknown and have to be estimated from the sample. The mean of the functions is estimated by

$$\widehat{\mu}(t) = \frac{1}{n} \sum_{k=1}^n l_k(t), \quad (16)$$

where  $n$  is the sample size. We estimate the kernel operator by

$$(\widehat{K}\phi)(s) = \int_{\mathcal{T}} \widehat{C}(s, t)\phi(t)dt, \quad (17)$$

where

$$\widehat{C}(s, t) = \frac{1}{n} \sum_{k=1}^n \{l_k(s) - \widehat{\mu}(s)\}\{l_k(t) - \widehat{\mu}(t)\}. \quad (18)$$

Estimates of the eigenfunctions and scores are computed from (17) and denoted by  $\widehat{\phi}_i$  and  $\widehat{\alpha}_{ik}, i = 1, \dots, m$ . Hörmann and Kokoszka (2010) show that these estimates are  $\sqrt{n}$ -consistent for a large group of stationary functional time series, which includes in particular linear functional processes. To choose the number of principal component functions  $m$  in Equation (3.2) there exist various rules. Here, the number is chosen such that at least an a priori fixed amount of variation in the data is explained by the principal component functions.

### 3.4 Forecasting functional time series

As noted above, in Equation (14) only the principal component scores  $\alpha$  depend on time, while the common mean function  $\mu$  and the principal component functions  $\phi$  are time-invariant. Hence, the dynamics over time of the load curves are fully captured by the dynamics of the principal component scores. The time dynamics of the first  $m$  estimated principal component scores can be modeled using time series analysis (Aue et al. (2014)). The problem of modelling the infinite dimensional functional data objects reduces thus to modeling a multivariate time series.

In our analysis a vector autoregressive model including exogeneous variables (VARX) turned out to be suitable to capture the dynamics of the principal component scores. The VARX model of order  $p$  is given by

$$\widehat{\alpha}_k = \sum_{i=1}^p \Phi_i \widehat{\alpha}_{k-i} + \beta x_k + \eta_k \quad (19)$$

where  $\widehat{\alpha}_k$  is the vector of estimated principal component scores,  $\Phi_i$  is a coefficient matrix,  $x_k$  are exogenous variables and  $\eta_k$  is a white noise process (Lütkepohl 2005).

A forecast of the principal component scores directly yields a forecast of the load curve:

$$\widehat{\widehat{l}}_{K+h}(t) = \widehat{\mu}(t) + \sum_{i=1}^m \widehat{\alpha}_{i,K+h} \widehat{\phi}_i(t), \quad (20)$$

where  $\widehat{\alpha}_{i,K+h}$  denotes the  $h$ -step ahead forecast of the principal component scores at time  $K$  and  $\widehat{l}_{K+h}(t)$  is the  $h$ -step ahead forecast of the load curve. Although suitable here, in a more general setting forecast models for the principal component scores are not restricted to VAR models and can be replaced by many other time series models. The next algorithm summarizes the aforementioned steps:

---

**Algorithm**

---

1. For fixed  $\tau \in (0, 1)$  and each  $k = 1, \dots, K$ , use the data  $Y_{t,k}, t = 1, \dots, 96$  to estimate the  $k - th$  generalized quantile curve  $l_{\tau,k}(t)$  by LAWS
2. Fix  $m$  and apply FPCA to  $l_{\tau,1}, \dots, l_{\tau,K}$ :
  - (a) Compute the empirical mean  $\widehat{\mu}$ , FPCs  $\widehat{\phi}_i$  and FPC scores  $\widehat{\alpha}_k = (\widehat{\alpha}_{1,k}, \dots, \widehat{\alpha}_{m,k})^\top$ ,  $k = 1, \dots, K, i = 1, \dots, m$
3. For fixed  $h$ , use the  $m$ -variate time series of empirical FPC scores  $\widehat{\alpha} = (\widehat{\alpha}_1 \dots \widehat{\alpha}_m)^\top$  to obtain  $h$ -step ahead forecast  $\widehat{\alpha}_{K+h} = (\widehat{\alpha}_{K+h,1} \dots \widehat{\alpha}_{K+h,m})^\top$ 
  - (a) Use a multivariate time series model, e.g. VAR model
  - (b) Include information from exogenous variables
4. Use  $\widehat{\alpha}_{K+h}$  to compute  $h$ -step ahead forecast of  $l_{K+h}$  as

$$\widehat{l}_{K+h}(t) = \widehat{\mu}(t) + \sum_{i=1}^m \widehat{\alpha}_{i,K+h} \widehat{\phi}_i(t)$$


---

### 3.5 Simulation Study

In this section simulations are used for illustrating the performance of the proposed method for FPCA with generalized quantiles. We run the simulation for independent as well as autocorrelated functional observations to demonstrate robustness to temporal dependence. For comparison, we follow the simulation setup of Guo et al. (2013) and Tran et al. (2014), who both suggest alternative approaches for modeling functional tail event curves. The data  $Y_{t,k}, k = 1, \dots, K, j = 1, \dots, T$  is simulated from the model

$$Y_{j,k} = \mu(t_j) + \alpha_{1,k} f_1(t_j) + \alpha_{2,k} f_2(t_j) + \varepsilon_{k,j} \quad (21)$$

where  $t_j$  are equidistant sampling points in  $[0, 1]$  with  $t_j = j/T$ ,  $\mu(t) = 1 + t + \exp\{-(t - 0.6)^2/0.05\}$  is the mean function,  $f_1(t) = \sqrt{2} \sin(2\pi t)$  and  $f_2(t) = \sqrt{2} \cos(2\pi t)$  are the principal component functions and  $\alpha_{1,k}$  and  $\alpha_{2,k}$  are principal component scores. The principal component scores are generated either (1) independently from a  $N(0, 36)$  and  $N(0, 9)$  distribution, respectively or (2) from a VAR(1) process with

$$\Phi_1 = \begin{pmatrix} -0.5 & -0.2 \\ 0.2 & 0.5 \end{pmatrix}.$$

The error  $\varepsilon_{k,t}$  is generated from three different distributions as specified in Table (2), where the first one is a light-tailed distribution, the second one is heavy-tailed and the third one exhibits heteroscedasticity. The simulation is run 200 times with two different setups: (1)  $T = 100$  grid points per curve and  $K = 20$  curves and (2)  $T = 150$  data points per curve and  $K = 50$  curves. Summary statistics of the mean squared errors (MSE) and the average run time in seconds of the simulations are given in Table (2). The magnitude of the average MSE does not differ substantially between the independent and the autocorrelated case. This confirms that the quality of the proposed methodology is not sensitive to temporal dependence between functional observations. The methodology performs worst for the fat tailed distribution, but well handles heteroscedasticity. Overall, the results are at least as good in terms of average MSE than those reported by Guo et al. (2013) and Tran et al. (2014).

		$N = 20, T = 100$		$N = 50, T = 150$	
		(1)	(2)	(1)	(2)
$\tau = 0.95$					
$\varepsilon \sim N(0, 0.5)$	Mean	0.0448	0.0400	0.0254	0.0233
	SD	0.0295	0.0271	0.0174	0.0170
	AT	3.4400	3.7200	10.1300	11.0100
$\varepsilon \sim t(5)$	Mean	0.2444	0.2290	0.1465	0.1428
	SD	0.2565	0.2396	0.1590	0.1729
	AT	3.7500	4.2300	10.3700	12.2500
$\varepsilon \sim N(0, \mu(t)0.5)$	Mean	0.0564	0.0518	0.0416	0.0500
	SD	0.0381	0.0340	0.0286	0.0389
	AT	3.5100	3.8200	10.2100	11.8000
$\tau = 0.05$					
$\varepsilon \sim N(0, 0.5)$	Mean	0.0433	0.0407	0.0259	0.0234
	SD	0.0285	0.0281	0.0177	0.0168
	AT	3.3900	3.7200	9.8000	10.3200
$\varepsilon \sim t(5)$	Mean	0.2447	0.2242	0.1480	0.1401
	SD	0.2508	0.2407	0.1644	0.1571
	AT	3.7100	4.6600	11.6200	12.0700
$\varepsilon \sim N(0, \mu(t)0.5)$	Mean	0.0521	0.0518	0.0499	0.0501
	SD	0.0379	0.0354	0.0385	0.0393
	AT	3.2600	3.7700	9.8700	12.2000

Table 2: Mean and standard deviation (SD) of MSE and average run time in seconds (AT) based on  $K = 200$  simulation runs

## 4. MODELING DYNAMICS OF ELECTRICITY DEMAND

As noted above the generalized quantile functions fully characterize the distribution of electricity demand and they are an alternative to modeling the distribution function directly. We model the dynamics of the 1%, 5%, 25%, 50%, 75%, 95% and 99% expectile functions using the described functional data approach for the two year insample period from 20100101 to 20113112. The expectile functions are obtained based on the LAWS algorithm with penalty terms  $\lambda_x$  and  $\lambda_\tau$  estimated by asymmetric cross validation.

The number of principal components is chosen such that at least 95% of the variation in the data is explained. Tables (3) and (4) show the explained variation of each principal component for all considered expectile levels  $\tau$  for the BU Saarbrücken and the TSO Amprion, respectively. With four principal components slightly more than 95% of the variation is captured. For all expectile levels the first PC is the dominant factor, explaining roughly three quarter of the variation in the curves.

PC	Expectile level $\tau$						
	1%	5%	25%	50%	75%	95%	99%
1	0.7221	0.7223	0.7232	0.7195	0.7247	0.7129	0.7233
2	0.1209	0.1213	0.1215	0.1220	0.1199	0.1241	0.1227
3	0.0695	0.0690	0.0687	0.0687	0.0698	0.0691	0.0722
4	0.0394	0.0409	0.0404	0.0420	0.0416	0.0410	0.0409
Total	0.9521	0.9536	0.9540	0.9524	0.9562	0.9572	0.9593

Table 3: Explained variance of the first four principal components for the BU Saarbrücken

PC	Expectile level $\tau$						
	1%	5%	25%	50%	75%	95%	99%
1	0.7985	0.7942	0.8007	0.7946	0.7952	0.7956	0.7892
2	0.0878	0.0893	0.0864	0.0883	0.0892	0.0887	0.0920
3	0.0583	0.0586	0.0553	0.0559	0.0547	0.0553	0.0541
4	0.0222	0.0237	0.0240	0.0255	0.0253	0.0252	0.0261
Total	0.9670	0.9659	0.9667	0.9644	0.9645	0.9650	0.9615

Table 4: Explained variance of the first four principal components for the TSO Amprion

Figure (3) shows the first four principal components for TSO and the BU. For both data sets the extracted PCs exhibit similar shapes. The first principal component reflects variations in the level of electricity load. A positive score on the first

principal component implies above average consumption, a negative score below average consumption (see Figure (4) for the scores). The second and third principal component capture variations in the height and location of peak load. The steepness of the load curves is modeled by the fourth principal component.

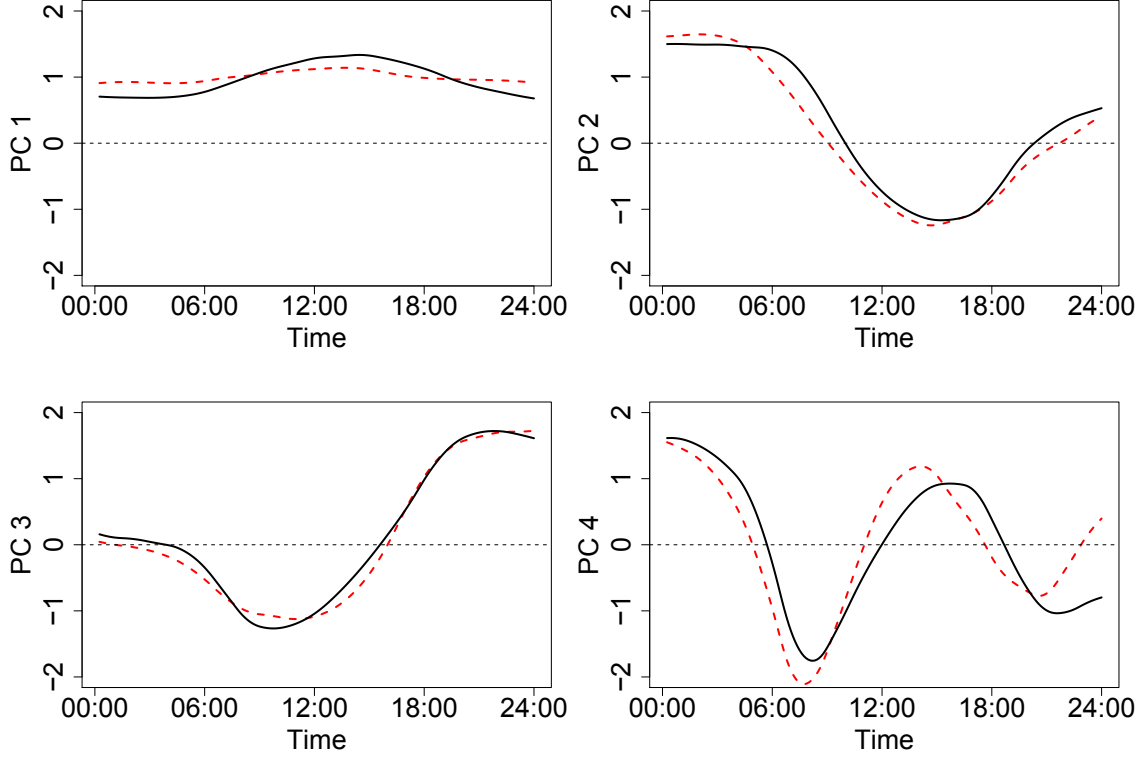


Figure 3: First four principal components corresponding to  $\tau = 50\%$  of the TSO Amprion (dashed) and the BU (solid)

The load of the TSO is the aggregated load of all BUs located in their area of responsibility. We compare the dynamics of the principal component scores for  $\tau = 0.5$  of the TSO and of the BU Saarbrücken in order to investigate to which degree the quantity risk in the BU is driven by the same factors as the overall market. The scores of the first factor of the TSO and the BU have a correlation of only 0.3 indicating that the level of electricity load of both is only moderately correlated. The correlation of the other scores and different levels of  $\tau$  is of similar magnitude.

The dynamics of the principal component scores are modeled by a VARX( $p$ ) model. The lag order  $p$  is chosen by the Akaike information criterion and equals  $p = 7$  for all expectile levels. Table (5) shows the estimation results of the exogenous variables for the TSO. Note that the values refer to deseasonalized data.

Both, daily average temperature and hours of sunshine have a strong negative effect on all four principal component scores, implying that they do not only affect the amount of electricity used, but also the allocation over the day. As expected, an increase in temperature results in lower electricity consumption, since in the area of

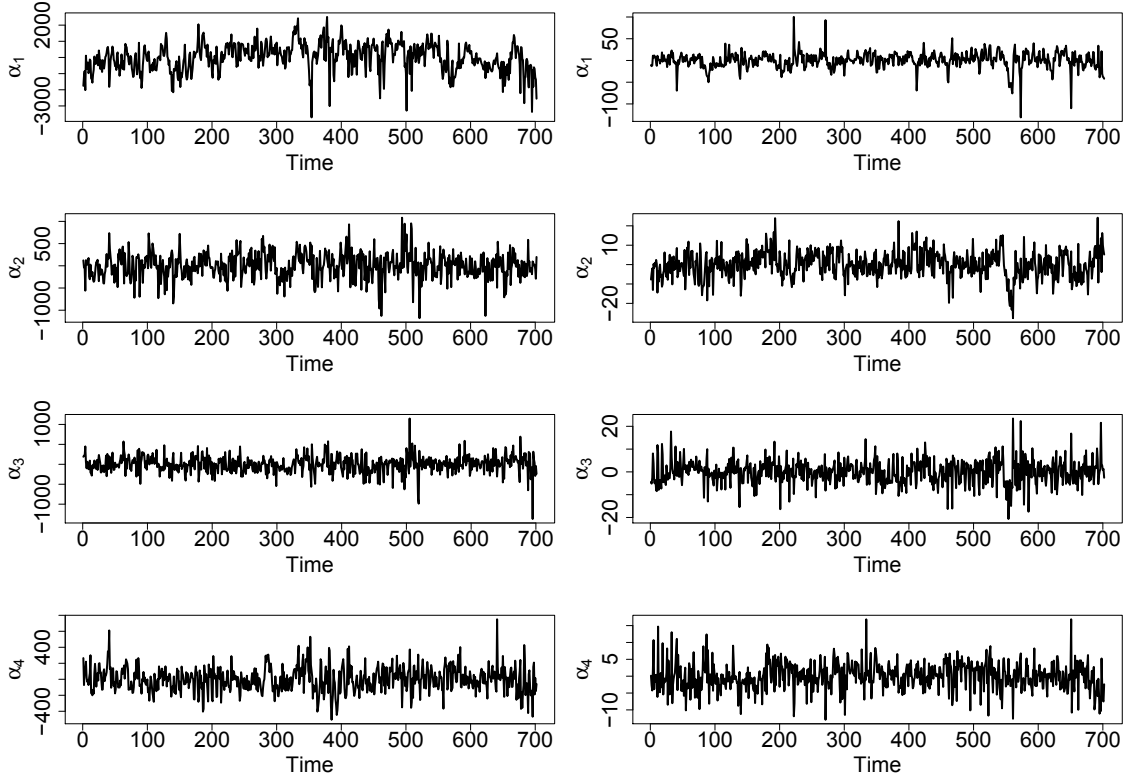


Figure 4: The scores corresponding to the first four principal components and  $\tau = 50\%$  of the TSO Amprion (left) and the BU Saarbrücken (right)

the TSO Amprion there is a high commonness of electric heating in form of night-storage heaters. Similarly, more hours of sunshine decrease electricity demand, for instance due to less demand for lighting and in-house consumption of solar energy. Peak demand is affected differently by temperature and hours of sunshine as reflected in the sign of the coefficients. While higher temperature causes more pronounced peaks, more hours of sunshine flatten the load curve.

For the BU (Table (6)) temperature does not have any effect on electricity load. This may be explained by the fact, that the share of electric heating is much lower in Saarbrücken than in the total area of the TSO. Therefore, demand is less sensitive to temperature.

In order to test the validity of our model, a multivariate portmanteau test (Hosking 1980) is performed, which suggests to reject the overall significance of residual autocorrelation for lag orders up to 50.

## 5. FORECAST EVALUATION

In this section we conduct an out-of sample forecast evaluation for the third year of electricity consumption data (2012) for the BU and the TSO.

	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$
$\tau = 5\%$				
Temperature	-22.314*** (7.175)	-14.805*** (2.903)	-11.665*** (2.620)	-3.629** (1.564)
Sunshine	-24.912*** (8.895)	38.058*** (3.599)	20.184*** (3.248)	-6.308*** (1.939)
$\tau = 50\%$				
Temperature	-22.156*** (7.141)	-14.708*** (2.896)	-11.276*** (2.591)	-3.940** (1.576)
Sunshine	-24.563*** (8.834)	38.198*** (3.582)	20.208*** (3.205)	-6.059*** (1.949)
$\tau = 95\%$				
Temperature	-21.892*** (7.187)	-13.771*** (2.930)	-11.587*** (2.593)	-3.790** (1.557)
Sunshine	-25.244*** (8.877)	37.850*** (3.619)	21.095*** (3.202)	-5.633*** (1.923)
<i>Note:</i>			*p<0.1; **p<0.05; ***p<0.01	

Table 5: Estimation Results TSO Amprion. Standard deviation in parenthesis.

	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$
$\tau = 5\%$				
Temperature	-0.038 (0.154)	-0.039 (0.058)	-0.016 (0.054)	-0.039 (0.041)
Sunshine	-0.950*** (0.191)	0.633*** (0.072)	0.194*** (0.067)	-0.204*** (0.051)
$\tau = 50\%$				
Temperature	-0.039 (0.154)	-0.037 (0.058)	-0.017 (0.054)	-0.041 (0.041)
Sunshine	-0.958*** (0.190)	0.631*** (0.072)	0.210*** (0.067)	-0.200*** (0.051)
$\tau = 95\%$				
Temperature	-0.048 (0.154)	-0.036 (0.058)	-0.022 (0.055)	-0.043 (0.041)
Sunshine	-0.978*** (0.190)	0.627*** (0.072)	0.228*** (0.068)	-0.171*** (0.050)
<i>Note:</i>			*p<0.1; **p<0.05; ***p<0.01	

Table 6: Estimation Results BU Saarbrücken. Standard deviation in parenthesis ().



## 5.1 Prediction comparison methods

In order to assess the performance of our forecasting approach we compare it to the results of three different benchmark models. As a first benchmark we choose the simple Deterministic Seasonal Component (DSC) (Eq. 2). The deterministic seasonal component is straightforward to estimate and in modified forms frequently used in industry. As a further benchmark we use forecasts that are provided by the TSO Amprion for their area of responsibility. Additionally, we compare our model to the triple seasonal Holt-Winter exponential smoothing (TSHW) model, which was proposed and applied to short-term load forecasting by Taylor (2010). The latter model accomodates yearly, weekly and daily seasonal cycles and has the advantage that no specification of the functional form is required. Furthermore, Taylor (2010) shows in an application to data from the UK that the TSHW model outperforms simpler exponential smoothing models and performs at least as well as seasonal ARMA models. The formulation of the TSHW model is given by

$$l_t = \alpha(y_t - d_{t-s_1} - w_{t-s_2} - a_{t-s_3}) + (1 - \alpha)l_{t-1} \quad (22)$$

$$d_t = \delta(y_t - l_t - w_{t-s_2} - a_{t-s_3}) + (1 - \delta)d_{t-s_1} \quad (23)$$

$$w_t = \omega(y_t - l_t - d_{t-s_1} - a_{t-s_3}) + (1 - \omega)w_{t-s_2} \quad (24)$$

$$a_t = \lambda(y_t - l_t - d_{t-s_1} - w_{t-s_2}) + (1 - \lambda)a_{t-s_3} \quad (25)$$

$$\hat{y}_{t+h} = l_t + d_{t-s_1+h} + w_{t-s_2+h} + a_{t-s_3+h} + \phi^k(y_t - l_{t-1} - d_{t-s_1} - w_{t-s_2} - a_{t-s_3}) \quad (26)$$

where  $l_t$  is the smoothed load,  $d_t, w_t$  and  $a_t$  are seasonal indices for daily, weekly and annual cycles and  $\alpha, \delta, \omega, \lambda$  are smoothing parameters.  $\hat{y}_{t+h}$  denotes the  $h$ -step ahead forecast at time  $t$  and the term including  $\phi$  is an adjustment for first-order autocorrelation.

## 5.2 Forecast Results

We evaluate the performance of the proposed forecasting methodology based on the root mean squared error (RMSE) defined as

$$RMSE_h = \sqrt{\frac{1}{96} \sum_{t=1}^{96} \{l_h(t) - \hat{l}_h(t)\}^2}$$

and the mean absolute percentage error (MAPE) given by

$$MAPE_h = \frac{1}{96} \sum_{t=1}^{96} \left| \frac{l_h(t) - \hat{l}_h(t)}{l_h(t)} \right|,$$

where  $h = 1, \dots, H$  denotes the day-ahead forecasts. Table (7) reports the mean of the RMSE and MAPE based on  $H = 351$  out-of-sample day-ahead forecasts over the year 2012. Meteorological forecasts of the explanatory variables are used as covariates. The proposed methodology clearly outperforms the benchmark models in terms of the given measures. Surprisingly, also the simple DSC performs relatively good and yields results similar to those of the TSHW exponential smoothing model.

	FDA	DSC	TSHW	TSO
	TSO			
RMSE	711.72	1021.16	932.81	3064.97
MAPE	0.027	0.045	0.043	0.13
	Balancing Unit			
RMSE	23.16	30.30	34.71	–
MAPE	0.043	0.062	0.066	–

Table 7: Mean of RMSE and MAPE for the TSO and the BU of functional data approach (FDA) with  $\tau = 0.5$ , deterministic seasonal component (DSC), triple seasonal Holt-Winter exponential smoothing (TSHW) and forecast provided by the transmission system operator (TSO).

This underlines the very regular structure in intra-daily electricity consumption. Figure (5) shows boxplots of the MAPE. The distribution of MAPE of the FDA is most narrow, however it contains several extreme values at the right tail. Overall, the figures indicate that our model performs quite well in point forecasting and results in forecasts with an average absolute deviation of only 2.7% for the TSO and 4.3% for the BU.

In order to evaluate the performance of the distributional forecasts, we look at the accuracy of the expectile estimates for  $\tau = 0.01, 0.05, 0.25, 0.75, 0.95, 0.99$ . As pointed out by Guler et al. (2014) a performance measure should be used that reflects the asymmetric loss function of expectiles. We choose the root mean weighted squared error (RMWSE) defined as

$$RMWSE_h^\tau = \sqrt{\frac{1}{96} \sum_{t=1}^{96} |\tau - \mathbf{I}_{\{l_h(t) < \widehat{l}_{\tau,h}(t)\}}| \{l_h(t) - \widehat{l}_{\tau,h}(t)\}^2}, \quad (27)$$

where  $\widehat{l}_\tau$  denotes the forecasted  $\tau$ -expectile curve. Summary statistics of the RMWSE for the different expectile levels based on  $H = 351$  day-ahead forecasts are given in Table (8). The results indicate high precision of the estimated tail curves. Figure (6) shows a plot of the forecasted mean together with forecast intervals corresponding to various expectile levels and the observed load for the TSO Amprion. The intervals spanned by the expectile curves provide information about the dispersion and asymmetry of next day's load. Furthermore, as pointed out above, quantiles and expectiles can be optimal point forecasts if the forecasters loss function is asymmetric. In energy demand forecasting, asymmetric loss functions occur if the cost for positive and negative imbalance is asymmetric, as it is the case in Scandinavian countries (Linnet 2005). In that case trading strategies based on quantiles and expectiles can be beneficial.

	Expectile level $\tau$						
	1%	5%	25%	50%	75%	95%	99%
TSO							
Mean	566.83	685.42	700.51	711.72	673.97	611.88	496.68
SD	440.98	494.81	420.20	416.40	391.32	432.82	373.50
BU							
Mean	13.93	16.82	20.90	23.16	24.47	24.20	23.36
SD	6.57	6.71	6.49	7.61	9.31	10.93	11.41

Table 8: Summary Statistics of RMWSE and MAWPE

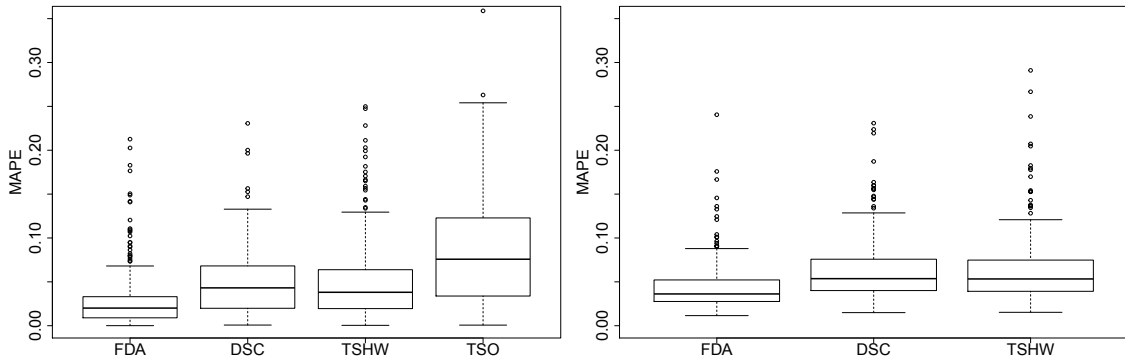


Figure 5: Boxplot of the mean absolute percentage forecasting error for the TSO (left) and the BU (right) of the functional data approach (FDA), the deterministic seasonal component (DSC), the triple seasonal Holt-Winter exponential smoothing model (TSHW) and the forecast provided by the transmission system operator (TSO).

## 6. CONCLUSION

In this article we show how to get powerful probabilistic short-term forecasts of intradaily electricity load employing functional data analysis methods with generalized quantile regression. Probabilistic forecasts yield important information about the uncertainty of future demand and are crucial for sustainable operation of electric utilities and for traders in the electricity market. This novel approach has several advantages: It allows for flexible inclusion of explanatory variables and does not require distributional assumptions for neither the tails nor the functional form of the tail time varying curves. The empirical analysis is based on quarter-hourly data from a TSO and a BU in western Germany. The proposed methodology identifies the main risk drivers of the electricity load by a number of factors: variations in the level of electricity load, variation in the height and location of peak load and the steepness of the load curves. We show how these factors are then to be (cor)related with climate effects.

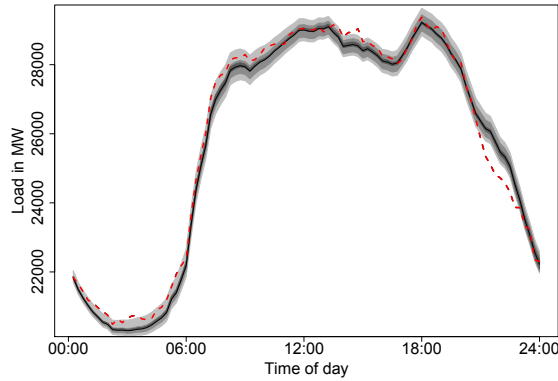


Figure 6: Forecasted expected load (black solid line) together with forecasts of  $\tau = 0.01, 0.05, 0.25, 0.75, 0.95, 0.99$  expectiles (gray shades) and observed load (red dashed line) for TSO Amprion on 20120125.

In a forecast comparison study we find that our methodology outperforms forecasts (a MAPE of 2.7%) provided by the TSO as well as those of the benchmark models.

## References

- Antoch, J., Prchal, L., De Rosa, M. R., and Sarda, P. (2008), “Functional linear regression with functional response: application to prediction of electricity consumption,” in *Functional and Operatorial Statistics*, Springer, pp. 23–29.
- Aue, A., Norinho, D. D., and Hörmann, S. (2014), “On the prediction of stationary functional time series,” *Journal of the American Statistical Association*.
- Bellini, F., Klar, B., Müller, A., and Rosazza-Gianin, E. (2014), “Generalized quantiles as risk measures,” *Insurance: Mathematics and Economics*, 54, 41–48.
- Bremnes, J. B. (2004), “Probabilistic wind power forecasts using local quantile regression,” *Wind Energy*, 7, 47–54.
- Cho, H., Goude, Y., Brossat, X., and Yao, Q. (2013), “Modeling and Forecasting Daily Electricity Load Curves: A Hybrid Approach,” *Journal of the American Statistical Association*, 108, 7–21.
- Cottet, R. and Smith, M. (2003), “Bayesian Modeling and Forecasting of Intraday Electricity Load,” *Journal of the American Statistical Association*, 98, 839–849.
- Engel, R. F., Granger, C. W. J., Rice, J., and Weiss, A. (1986), “Semiparametric Estimated of the Relation Between Weather and Electricity sales,” *Journal of the American Statistical Association*, 81, 310–320.
- Eurostat (2014), “Supply, transformation, consumption - electricity - annual data,” <http://epp.eurostat.ec.europa.eu>, accessed: 2014-03-17.

- Ferraty, F. and Vieu, P. (2006), *Nonparametric functional data analysis: theory and practice*, Springer.
- Gneiting, T. (2011), “Quantiles as optimal point forecasts,” *International Journal of Forecasting*, 27, 197–207.
- Goia, A., May, C., and Fusai, G. (2010), “Functional clustering and linear regression for peak load forecasting,” *International Journal of Forecasting*, 26, 700–711.
- Guler, K., Ng, P. T., and Xiao, Z. (2014), “Mincer-Zarnovitz Quantile and Expectile Regressions for Forecast Evaluations under Asymmetric Loss Functions,” *North-ern Arizona University, Working Paper Series*.
- Guo, M., Zhou, L., Härdle, W. K., and Huang, J. Z. (2013), “Functional Data Analysis of Generalized Regression Quantiles,” *Statistics and Computing*.
- Harvey, A. and Koopman, Siem, J. (1993), “Forecasting Hourly Electricity Demand Using Time-Varying Splines,” *Journal of the American Statistical Association*, 1228–1236.
- Hörmann, S. and Kokoszka, P. (2010), “Weakly dependent functional data,” *The Annals of Statistics*, 38, 1845–1884.
- Hosking, J. R. (1980), “The multivariate portmanteau statistic,” *Journal of the American Statistical Association*, 75, 602–608.
- Hyndman, R. J. and Fan, S. (2010), “Density forecasting for long-term peak electricity demand,” *Power Systems, IEEE Transactions on*, 25, 1142–1153.
- Koenker, R. (2005), *Quantile regression*, Cambridge University Press.
- Koenker, R. and Bassett Jr, G. (1978), “Regression quantiles,” *Econometrica: Journal of the Econometric Society*, 46, 33–50.
- Leutbecher, M. and Palmer, T. (2008), “Ensemble forecasting,” *Journal of Computational Physics*, 227, 3515–3539.
- Linnet, U. (2005), “Tools supporting wind energy trade in deregulated markets,” Ph.D. thesis, Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark.
- Lütkepohl, H. (2005), *New introduction to multiple time series analysis*, Springer.
- Newey, W. K. and Powell, J. L. (1987), “Asymmetric least squares estimation and testing,” *Econometrica: Journal of the Econometric Society*, 55, 819–847.
- Pinson, P., Chevallier, C., and Kariniotakis, G. N. (2007), “Trading wind generation from short-term probabilistic forecasts of wind power,” *Power Systems, IEEE Transactions on*, 22, 1148–1156.
- Ramsay, J. O. and Silverman, B. W. (2002), *Applied functional data analysis: methods and case studies*, Springer.

- (2005), *Functional data analysis*, Springer.
- Schnabel, S. (2011), “Expectile smoothing: new perspectives on asymmetric least squares. An application to life expectancy,” Ph.D. thesis, Utrecht University.
- Schnabel, S. K. and Eilers, P. H. C. (2013), “Simultaneous estimation of quantile curves using quantile sheets,” *AStA Advances in Statistical Analysis*, 97, 77–87.
- Shang, H. L. (2013), “Functional time series approach for forecasting very short-term electricity demand,” *Journal of Applied Statistics*, 40, 152–168.
- (2014), “A survey of functional principal component analysis,” *AStA Advances in Statistical Analysis*, 98, 121–142.
- Tay, A. S. and Wallis, K. F. (2000), “Density Forecasting: A Survey,” *Journal of Forecasting*, 19, 235–254.
- Taylor, J. W. (2008), “Estimating Value at Risk and Expected Shortfall Using Expectiles,” *Journal of Financial Econometrics*, 6, 231–252.
- (2010), “Triple seasonal methods for short-term electricity demand forecasting,” *European Journal of Operational Research*, 204, 139–152.
- Taylor, J. W. and Buizza, R. (2002), “Neural Network Load Forecasting with Weather Ensemble Predictions,” *Power Systems, IEEE Transactions on*, 17, 626–632.
- Taylor, J. W. and McSharry, P. E. (2007), “Short-term load forecasting methods: An evaluation based on European data,” *Power Systems, IEEE Transactions on*, 22, 2213–2219.
- Tran, N. M., Osipenko, M., and Härdle, W. K. (2014), “Principal Component Analysis in an Asymmetric Norm,” *Humboldt University Berlin, CRC 649 Discussion Paper*.
- Weron, R. (2007), *Modeling and forecasting electricity loads and prices: A statistical approach*, Wiley.

# SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Principal Component Analysis in an Asymmetric Norm" by Ngoc Mai Tran, Maria Osipenko and Wolfgang Karl Härdle, January 2014.
- 002 "A Simultaneous Confidence Corridor for Varying Coefficient Regression with Sparse Functional Data" by Lijie Gu, Li Wang, Wolfgang Karl Härdle and Lijian Yang, January 2014.
- 003 "An Extended Single Index Model with Missing Response at Random" by Qihua Wang, Tao Zhang, Wolfgang Karl Härdle, January 2014.
- 004 "Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey" by Helmut Lütkepohl, January 2014.
- 005 "Functional stable limit theorems for efficient spectral cointegration estimators" by Randolf Altmeyer and Markus Bibinger, January 2014.
- 006 "A consistent two-factor model for pricing temperature derivatives" by Andreas Groll, Brenda López-Cabrera and Thilo Meyer-Brandis, January 2014.
- 007 "Confidence Bands for Impulse Responses: Bonferroni versus Wald" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, January 2014.
- 008 "Simultaneous Confidence Corridors and Variable Selection for Generalized Additive Models" by Shuzhuan Zheng, Rong Liu, Lijian Yang and Wolfgang Karl Härdle, January 2014.
- 009 "Structural Vector Autoregressions: Checking Identifying Long-run Restrictions via Heteroskedasticity" by Helmut Lütkepohl and Anton Velinov, January 2014.
- 010 "Efficient Iterative Maximum Likelihood Estimation of High-Parameterized Time Series Models" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, January 2014.
- 011 "Fiscal Devaluation in a Monetary Union" by Philipp Engler, Giovanni Ganelli, Juha Tervala and Simon Voigts, January 2014.
- 012 "Nonparametric Estimates for Conditional Quantiles of Time Series" by Jürgen Franke, Peter Mwita and Weining Wang, January 2014.
- 013 "Product Market Deregulation and Employment Outcomes: Evidence from the German Retail Sector" by Charlotte Senftleben-König, January 2014.
- 014 "Estimation procedures for exchangeable Marshall copulas with hydrological application" by Fabrizio Durante and Ostap Okhrin, January 2014.
- 015 "Ladislaus von Bortkiewicz - statistician, economist, and a European intellectual" by Wolfgang Karl Härdle and Annette B. Vogt, February 2014.
- 016 "An Application of Principal Component Analysis on Multivariate Time-Stationary Spatio-Temporal Data" by Stephan Stahlschmidt, Wolfgang Karl Härdle and Helmut Thome, February 2014.
- 017 "The composition of government spending and the multiplier at the Zero Lower Bound" by Julien Albertini, Arthur Poirier and Jordan Roulleau-Pasdeloup, February 2014.
- 018 "Interacting Product and Labor Market Regulation and the Impact of Immigration on Native Wages" by Susanne Prantl and Alexandra Spitz-Oener, February 2014.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



# SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 019 "Unemployment benefits extensions at the zero lower bound on nominal interest rate" by Julien Albertini and Arthur Poirier, February 2014.
- 020 "Modelling spatio-temporal variability of temperature" by Xiaofeng Cao, Ostap Okhrin, Martin Odening and Matthias Ritter, February 2014.
- 021 "Do Maternal Health Problems Influence Child's Worrying Status? Evidence from British Cohort Study" by Xianhua Dai, Wolfgang Karl Härdle and Keming Yu, February 2014.
- 022 "Nonparametric Test for a Constant Beta over a Fixed Time Interval" by Markus Reiß, Viktor Todorov and George Tauchen, February 2014.
- 023 "Inflation Expectations Spillovers between the United States and Euro Area" by Aleksei Netšunajev and Lars Winkelmann, March 2014.
- 024 "Peer Effects and Students' Self-Control" by Berno Buechel, Lydia Mechtenberg and Julia Petersen, April 2014.
- 025 "Is there a demand for multi-year crop insurance?" by Maria Osipenko, Zhiwei Shen and Martin Odening, April 2014.
- 026 "Credit Risk Calibration based on CDS Spreads" by Shih-Kang Chao, Wolfgang Karl Härdle and Hien Pham-Thu, May 2014.
- 027 "Stale Forward Guidance" by Gunda-Alexandra Detmers and Dieter Nautz, May 2014.
- 028 "Confidence Corridors for Multivariate Generalized Quantile Regression" by Shih-Kang Chao, Katharina Proksch, Holger Dette and Wolfgang Härdle, May 2014.
- 029 "Information Risk, Market Stress and Institutional Herding in Financial Markets: New Evidence Through the Lens of a Simulated Model" by Christopher Boortz, Stephanie Kremer, Simon Jurkatis and Dieter Nautz, May 2014.
- 030 "Forecasting Generalized Quantiles of Electricity Demand: A Functional Data Approach" by Brenda López Cabrera and Franziska Schulz, May 2014.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

